# The clustering of sub-mJy radio sources in The Bootes Deep Field

R.J. Wilman, H.J.A. Röttgering, R.A. Overzier, and M.J. Jarvis

Sterrewacht Leiden, Postbus 9513, 2300 RA Leiden, The Netherlands.

4 November 2002

**ABSTRACT**

We measure the angular clustering of $\sim 2000$ radio sources in The Bootes Deep Field, covering 5.3 deg$^2$ down to $S_{1.4\mathrm{GHz}} = 0.2$ mJy. With reference to work by Blake & Wall, we show that the size distribution of multi-component radio galaxies dominates the overall clustering signal, and that its amplitude extrapolates smoothly from their measurements above 5 mJy. The upper limits on any true galaxy-galaxy clustering are consistent with the clustering of sub- mJy radio-loud AGN being effectively diluted by the more weakly-clustered *IRAS*-type starburst galaxies. Source count models imply that the survey contains $\simeq 400$ of the latter galaxies above 0.2 mJy out to $z \sim 1-2$. Measurement of their clustering must await their identification via the optical and infrared data due on this field.

**Key words:** surveys – galaxies:active – galaxies:starburst – large-scale structure of Universe

## 1 INTRODUCTION

Measurements of the clustering of galaxies can shed light on the formation of large scale structure and of the galaxies within it. In the local Universe ($z < 0.1$), optically-selected galaxies are unbiased tracers of the mass, following a spatial correlation function of the form $\xi(r) = (r/r_0)^{-\gamma}$, with $r_0 = 5.4h^{-1}$ Mpc and $\gamma = 1.8$ (Davis & Peebles 1983). Results from the 2dF QSO redshift survey show that at $\bar{z} = 1.3$, QSOs are comparably clustered ($r_0 = 5.7h^{-1}$ Mpc; Shanks et al. 2001). In contrast, the local population of *IRAS* starburst galaxies are less strongly clustered ($r_0 \simeq 3.8h^{-1}$ Mpc; Saunders et al. 1992), whilst the opposite is true of early-type galaxies (for local $L \gtrsim L^\star$ ellipticals $r_0 = 9-11h^{-1}$ Mpc, e.g. Guzzo et al. 1997, Willmer et al. 1998).

Extending such work to higher redshift populations can constrain the galaxy systems into which they evolve: e.g. Adelberger (2000) measured the clustering of $z \sim 1$ Balmer-break galaxies and derived $r_0 = 3.0h^{-1}$ Mpc (comoving). This is consistent with the $z \sim 1$ Balmer-break galaxies being both unbiased tracers of mass and the progenitors of local $L^\star$ galaxies. The comparable comoving space densities of the two populations supports this interpretation. In contrast, at $z \sim 3$ the Lyman-break galaxies are more strongly clustered (see Porciani & Giavalisco 2002), implying that they are strongly biased relative to the mass distribution. Similarly, analysis of the extremely red objects (EROs) demonstrates that passively evolving early-type galaxies at $z \sim 1.2$ have $r_0 = 12 \pm 3h^{-1}$ Mpc, comparable to that of local $L^\star$ ellipticals (Daddi et al. 2001; McCarthy et al. 2001). This rules out a scenario in which Lyman-break galaxies evolve into local bright ellipticals, and requires that the bias evolves with redshift such that the comoving density of ellipticals is invariant to $z \sim 1$ and beyond (see also Moustakas & Somerville 2002).

In contrast, radio surveys probe galaxy clustering over larger scales, as they cover large areas of sky with a broad redshift distribution. This does, however, reduce the clustering signal when projected on the sky. Nevertheless, numerous positive measurements of the angular clustering of radio sources have been made. The most complete analysis to date is that of Blake & Wall (2002) (hereafter BW02), who measured the angular correlation function of sources in the 1.4 GHz NVSS from flux thresholds of 50 down to 5 mJy (see also Overzier et al. 2002). They found that the clustering at small angles is due to the size distribution of multi-component radio galaxies, and that the galaxy-galaxy correlation function has the universal slope of $\gamma = 1.8$, and (under certain assumptions) a correlation length of $r_0 \sim 6h^{-1}$ Mpc.

Below $\sim 1$ mJy (at 1.4 GHz) there is evidence for the appearance of a different radio source population which may replace the radio-loud AGN as the dominant population below a few 100$\mu$Jy (Windhorst 1984, Windhorst et al. 1985, Benn et al. 1993). These sources are more distant analogues of the dusty starburst galaxies selected by IRAS, and models of the radio source counts require that this population undergo substantial luminosity evolution out to $z \sim 1$ (Rowan-Robinson et al. 1993; hereafter RR93). Measurements of the clustering of the sub-mJy population could thus extend the local IRAS clustering measurements of starburst galaxies out to $z \sim 1$, for comparison with those of the Balmer-break galaxies, the EROs referred to above and the dusty, star-forming EROs at $z \sim 1$ which have $r_0$ no larger than $2.5h^{-1}$ Mpc (Daddi et al. 2002). Such a measurement was attempted by Georgakakis et al. (2000) for sources with $S_{1.4} > 0.5$ mJy in the $\simeq 3$deg$^2$ Phoenix radio survey (Hopkins et al. 1998). Whilst they found some evidence for clustering, the uncertainties were such that they could not determine whether the starbursts have a lower $r_0$ than the radio-loud AGN.

To overcome these limitations we have initiated The Bootes Deep Field, a radio survey covering $\simeq$6 deg$^2$ down to a limiting

$5\sigma$ sensitivity of 140 $\mu$Jy at 1.4 GHz (de Vries et al. 2002). The source finding algorithm identifies 3172 distinct sources, of which $\simeq 10$ per cent are resolved by the $13 \times 27$ arcsec beam (the actual number of sources used in our angular clustering analysis is smaller than this, due to our use of a subset of the data with a slightly higher flux limit, and the use of certain algorithms to identify multiple component sources; see section 2). It will be complemented by imaging in six optical and near-infrared wavebands (as part of the *NOAO Deep Wide-Field Survey*), and at longer infrared wavelengths with SIRTF. These data will enable morphological classification of the radio sources into starburst or AGN, photometric redshift estimation, and an examination of the far-infrared:radio correlation for starburst galaxies out to cosmological distances (building on work by Garrett 2002). Direct measurement of the real space correlation functions of both radio-loud AGN and starbursts (and of the cross-correlation between them) will also be possible to $z \sim 1$. Furthermore, these measurements will be free from the confusion caused by multi-component radio sources in existing catalogues without identifications. Until this dataset is assembled, however, we are confined to use of the angular correlation function, and it is these measurements and their interpretation that we present here.

## 2 ANGULAR CLUSTERING MEASUREMENTS

Full details of the Bootes survey characteristics and source extraction procedure are given in de Vries et al. (2002). As detailed therein, the survey is complete down to a limiting flux density of 0.2 mJy, as determined by the point at which the source counts exhibit the first systematic deviation from a low-order polynomial fit (see their Fig. 9; N.B. The source counts shown in de Vries et al. are all too low by a factor 2.035 due to a numerical error in the construction of this plot). Since extracted sources must have peak fluxes at least 5 times higher than the local rms noise level, for the clustering analysis we exclude the corners of the survey area where the latter exceeds 0.04 mJy; the remaining area covers 5.32 deg$^2$.

The angular two-point correlation function $\omega(\theta)$ is defined through the expression for the probability, $\delta P$, of finding two sources in solid angle elements $\delta\Omega_1$ and $\delta\Omega_2$, with angular separation $\theta$:

$$\delta P = N^2 \delta\Omega_1 \delta\Omega_2 (1 + \omega(\theta)), \quad (1)$$

where $N$ is the mean areal source density. The function $\omega(\theta)$ is therefore a measure of the deviation from a random distribution. It can be calculated by comparing the number of source pairs within a given range of angular separation with the number of pairs in a large random catalogue covering the same area. Of the variety of estimators for $\omega(\theta)$ which have been proposed, we follow Rengelink (1999) and use that due to Hamilton (1993) because of its robustness at large angular scales:

$$\omega(\theta) = 4 \frac{n_{\rm D} n_{\rm R}}{(n_{\rm D} - 1)(n_{\rm R} - 1)} \frac{DD \cdot RR}{DR \cdot DR} - 1 \quad (2)$$

where $DD$ is the number of data-data pairs within the angular bin centred on separation $\theta$, and $RR$ and $DR$ are the numbers of random-random and data-random pairs, respectively, within the same separation interval. We use a random catalogue containing $n_{\rm R} = 25000$ sources, vastly exceeding the size of the data catalogue ($n_{\rm D}$). Errors on the individual $\omega(\theta)$ points are computed with the method of Ling, Barrow & Frenk (1986), by calculating the standard deviation in $\omega(\theta)$ among 20 pseudo-random resamples of the observational dataset. Such estimates exceed the Poisson errors, which are correct only for unclustered data.

Using this procedure, $\omega(\theta)$ was computed in equally-spaced logarithmic angular intervals between 0.25 and 155 arcmin for several subsamples with lower limiting fluxes between 0.2 and 2 mJy. Each was then fitted with a function of the form:

$$\omega(\theta)_{\rm fit} = A\theta^{-\delta} - C \quad (3)$$

with $\theta$ in degrees. The quantity C (the integral constraint; Groth & Peebles 1977) is a bias resulting from the finite boundary of the survey and is given by:

$$C = \frac{1}{\Omega^2} \int\int \omega(\theta) d\Omega_1 d\Omega_2. \quad (4)$$

For the conventional power-law form $\omega(\theta) = A\theta^{-\delta}$, Monte-Carlo computation of the integral yields $C = 1.154A$ and $1.392A$ for $\delta = 0.8$ and 1.1, respectively.

A complication in the analysis of radio source clustering is the signal at small angular separations (below a few arcminutes) caused by sources with multiple components. This problem will be largely alleviated when such sources can be reliably identified with the optical and infrared data. In the meantime, we begin by adopting some prescriptions used by Georgakakis et al. (2000), and before that by Magliocchetti et al. (1998) and Cress et al. (1996), to identify the genuine multi-component sources prior to measuring the angular correlation function. Using the $\theta \propto \sqrt{S}$ relation found by Oort (1987), we consider as a single object all doubles with $\theta < 20\sqrt{F_{\rm total}}$, where $\theta$ is their separation in arcseconds and $F_{\rm total}$ their summed flux density in mJy; furthermore, we only collapse doubles whose individual component fluxes differ by less than a factor of 4 (since components of genuine doubles are expected to have correlated fluxes; Magliocchetti et al. 1998). Of the 214 double sources identified by the source extraction algorithm, 50 are identified as genuine by these criteria. The 48 sources with 3 or more components (39 triples, 9 quadruples) were examined by eye and subjective criteria (e.g. relative component fluxes, morphology) were used to decide whether or not to treat them as a single source, resulting in the assignment of 78 separate sources.

Fits to the data were performed between separations of 1.5 and 20 arcmins, and are shown in Fig. 1 for the 0.2 and 2 mJy subsamples. The fitted parameters for these and other subsamples are listed in Table 1. As expected, the amplitudes are lower when fitted with steeper power-laws. The variation of amplitude with flux limit for our survey is shown in Fig. 2. Although the error bars are large, our measured amplitudes of $\simeq 0.01$ for flux limits of 1–2 mJy are consistent with the results of Georgakakis et al (2000) for the Phoenix survey, and with those of Cress et al. (1996) who measured an amplitude of $\simeq 0.008$ for FIRST sources in the flux interval 1–2 mJy.

However, clustering analyses of the NVSS and FIRST radio surveys by BW02 and Overzier et al. (2002) are at odds with the earlier measurements from the FIRST survey. Both papers question the efficacy of the procedures previously used (which we also adopt) to identify the multi-component sources: they find that the angular correlation function below 6 arcmin is dominated by multi-component sources and that the true cosmological clustering amplitude is essentially constant at $\simeq 10^{-3}$ from 3 mJy to $\simeq 50$ mJy. It seems implausible that the amplitude could drop by a factor of 10 from flux limits of 1–2 mJy to 3 mJy. We thus conclude that the increase in amplitude from 0.2 to 2 mJy in Fig. 2 is due to the residual effects of multi-component sources at the higher fluxes. Indeed, Fig. 1 shows excess signal above the fitted power-law at small angles for the 2 mJy subsample.

**Table 1.** Fitted amplitudes of the angular correlation function after attempted removal of multi-component sources. The errors are $1\sigma$.

| Flux density ( mJy) | $10^3 A(\delta = 0.8)$ | $10^3 A(\delta = 1.1)$ |
|---|---|---|
| $> 0.2$ | $2.1 \pm 1.6$ | $0.86 \pm 0.67$ |
| $> 0.4$ | $2.0 \pm 3.0$ | $0.86 \pm 1.3$ |
| $> 0.6$ | $4.5 \pm 3.9$ | $1.9 \pm 1.6$ |
| $> 0.8$ | $4.2 \pm 5.4$ | $1.8 \pm 2.3$ |
| $> 1.0$ | $4.9 \pm 4.8$ | $2.3 \pm 2.0$ |
| $> 1.5$ | $9.4 \pm 7.7$ | $3.9 \pm 3.2$ |
| $> 2.0$ | $12 \pm 8.3$ | $5.4 \pm 3.6$ |

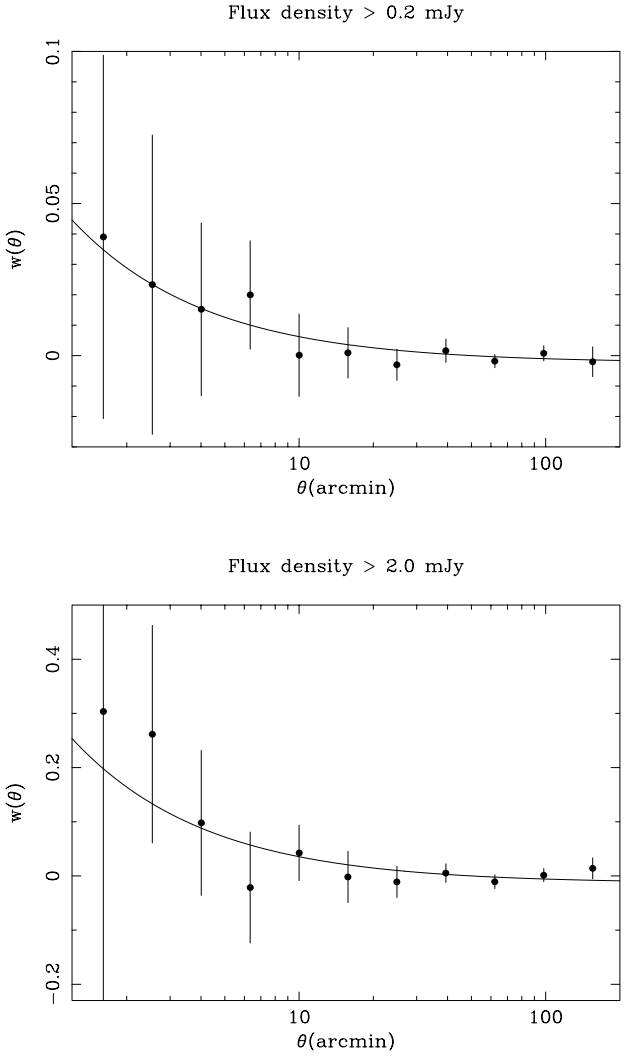


**Figure 1.** Angular correlation functions for subsamples with lower flux limits of 0.2 and 2 mJy, after the attempted removal of genuine multi-component radio sources, as described in section 2. The solid lines show fits to the data over $1.5 \leq \theta \leq 20$ arcmin with the function given in eqn. 3, for $\delta = 0.8$. There is an excess of data over the model at the smallest separations in the 2 mJy data, likely due to the residual effects of the multi-component sources.

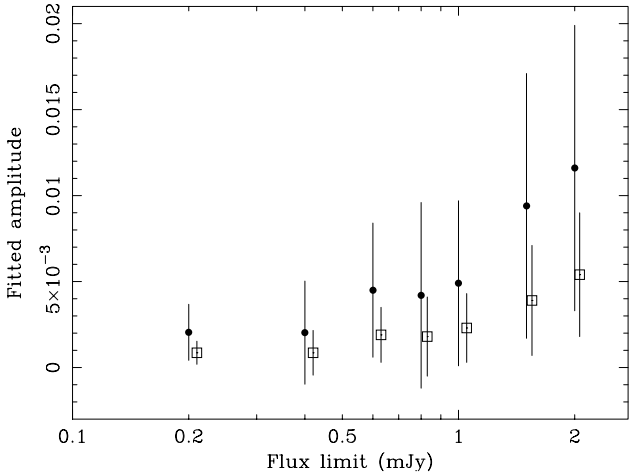


**Figure 2.** Variation of the fitted amplitude of the angular correlation function with flux limit in The Bootes Deep Field (after attempted removal of multi-component sources from the catalogue), for assumed power-law indices of $\delta = 0.8$ (filled circles) and $\delta = 1.1$ (open squares); for clarity the latter points have been displaced slightly in flux.

**Table 2.** Fitted amplitudes of the size-distribution and cosmological clustering power-laws derived from a two component fit. The errors are $1\sigma$.

| Flux density ( mJy) | size power-law $10^7 A(\delta = 3.4)$ | cosmological clustering $10^3 A(\delta = 0.8)$ |
|---|---|---|
| $> 0.2$ | $2.4 \pm 1.2$ | $0.55 \pm 1.4$ |
| $> 0.3$ | $1.9 \pm 1.6$ | $0.38 \pm 1.9$ |
| $> 0.4$ | $2.2 \pm 1.7$ | $1.2 \pm 2.5$ |
| $> 0.6$ | $4.1 \pm 2.1$ | $1.5 \pm 3.7$ |
| $> 0.8$ | $3.4 \pm 3.4$ | $2.1 \pm 4.4$ |
| $> 1.0$ | $4.8 \pm 5.3$ | $2.5 \pm 4.7$ |
| $> 1.5$ | $9.7 \pm 7.6$ | $4.8 \pm 6.9$ |
| $> 2.0$ | $14 \pm 8.3$ | $4.3 \pm 8.5$ |

In an attempt to overcome the effects of multi-component sources, we instead apply none of the above component combining procedures and fit the resulting correlation functions with the sum of two power-laws of the form shown in eqn (3). The first has $\delta = 3.4$ (fixed) and is due to the multi-component sources which dominate the signal below 6 arcmin, as found by BW02 and Overzier et al. (the corresponding integral constraint is $A = 1191C$, with the correlation function truncated below the angular resolution of the survey to keep the integral in eqn. 4 finite). The second has $\delta = 0.8$ (fixed) and is due to the cosmological clustering. Fits are performed to data above 1 arcmin, as shown in Fig. 3 for the 0.2 and 2 mJy sub- samples. Table 2 lists the results. There is now no significant detection of cosmological clustering and the effects of the multi-component sources dominate the $\omega(\theta)$ signal. Furthermore, Fig. 4 shows that the amplitude of the size distribution power-law is consistent with the extrapolation to fainter fluxes of the $1/\sigma$ dependence ($\sigma$ being the surface density of radio sources) found from 5 to 50 mJy by BW02. With reference to their section 3, this is consistent with the quantity $e/\bar{n}$ varying by no more than a factor of $\simeq 2$ from 50 to 0.2 mJy, as judged from the scatter around the fit in Fig. 4 ($e$ being the fraction of sources observed to have multiple components, $\bar{n}$ the average number of radio components per source).

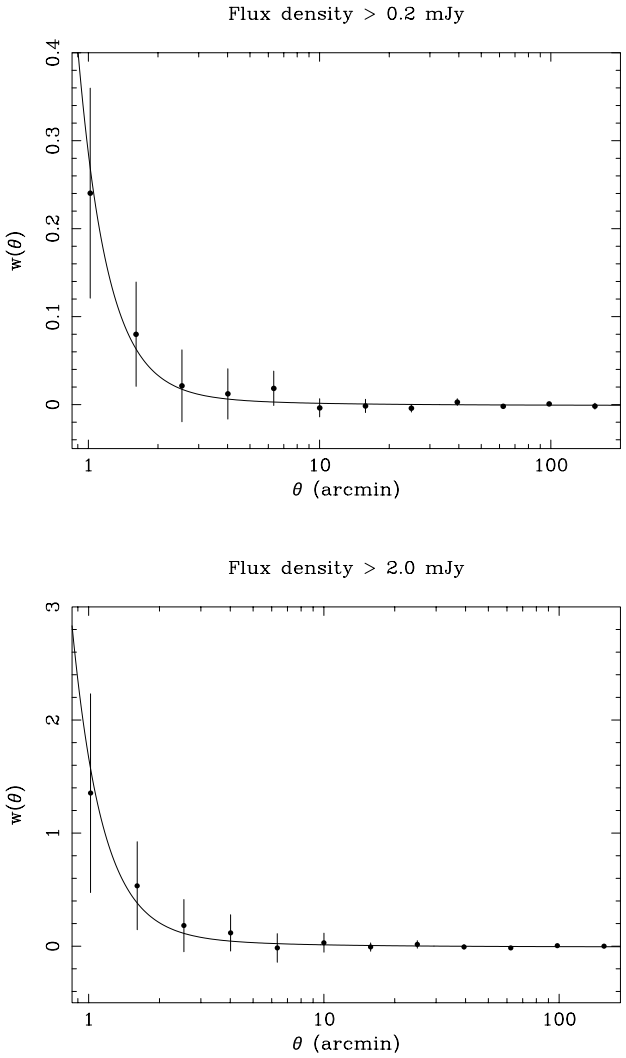


**Figure 3.** Angular correlation functions for subsamples with lower flux limits of 0.2 and 2 mJy, *without* applying any component combining procedures to the catalogue; the solid lines are fits to the data above 1 arcmin with a double power-law, as described in the the text. Note the different y-axis scales.

## 3 INTERPRETATION

We now model the evolution of the space density and clustering of the sub- mJy radio source population, to determine in particular the effect on the clustering signal of the appearance of the starburst galaxies. The spatial and angular correlation functions, $\xi(r)$ and $\omega(\theta)$, are related via the relativistic Limber equation (see e.g. Peebles 1980; Magliocchetti et al. 1999); when $\xi(r,z) = (r/r_0)^{-\gamma}(1+z)^{-(3+\epsilon)}$ ($r$ and $r_0$ being proper lengths), $w(\theta) = A\theta^{-(\gamma-1)}$ ($\theta$ in radians) where for a spatially flat cosmology with non-zero cosmological constant ($\Omega_\Lambda + \Omega_M = 1$):

$$A = B\frac{\int_0^\infty N^2(z)(1+z)^{\gamma-3-\epsilon}x^{1-\gamma}Q(z)dz}{\left(\int_0^\infty N(z)dz\right)^2} \quad (5)$$

with

$$B = \sqrt{\Omega_M}\left(\frac{r_0 H_0}{c}\right)^\gamma \frac{\Gamma(\frac{1}{2})\Gamma(\frac{\gamma}{2})}{\Gamma(\frac{\gamma-1}{2})}, \quad (6)$$

$$Q(z) = \left[(1+z)^3 + \Omega_M^{-1} - 1\right]^{0.5}, \quad (7)$$

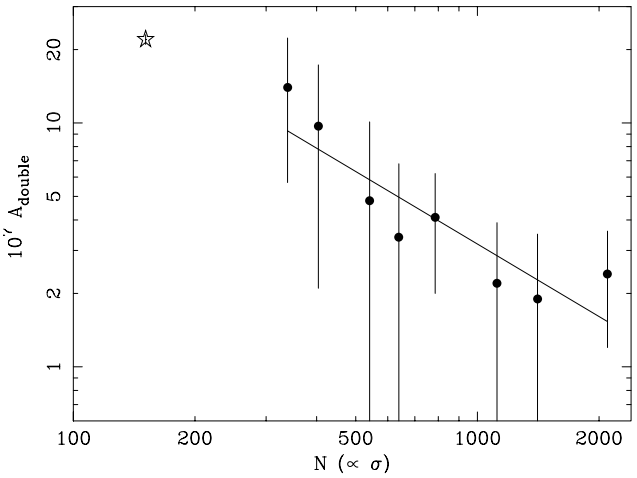


**Figure 4.** The circles represent the amplitude of the contribution to $\omega(\theta)$ from the multi-component sources, from Table 2. $N$ is the number of sources above the corresponding flux limit ($\propto \sigma$, the surface density – with no component-combining procedures applied). The fit to these points with $A_{\rm double} \propto N^\alpha$ is shown (with $\alpha = -0.99 \pm 0.6$), and is consistent with the $N^{-1}$ extrapolation from the higher flux limits found by BW02 (whose point for sources $> 5$ mJy is indicated by the star).

$$x = \frac{1}{\sqrt{\Omega_M}}\int_0^z \frac{dz}{Q(z)}, \quad (8)$$

and $N(z)$ is the redshift distribution of sources above the flux limit. When a sample comprises two sub-populations, A and B (in this case AGN and starbursts), with different clustering properties, the signal for the whole is given by:

$$\omega_{\rm eff} = f_A^2\omega_A + f_B^2\omega_B + 2f_A f_B \omega_{AB}, \quad (9)$$

where $f_A$ and $f_B$ are the fractions of the population in the two classes, and $\omega_A$ and $\omega_B$ their individual correlation functions. The term $\omega_{AB}$ is the cross-correlation between the two populations. In our analysis, we use the Limber equation to compute separate $\omega(\theta)$ for the AGN and starbursts, and then combine them using eqn. (9). For simplicity, we set the cross-correlation to zero.

Redshift distributions for each flux-limited subsample were computed using the Dunlop and Peacock (1990) pure luminosity evolution model for the combined population of flat ($\alpha = 0$; $S_\nu \propto \nu^\alpha$) and steep ($\alpha = -0.8$) spectrum AGN (taking the parameters from their table C3, shifted to 1.4 GHz). For the star-forming galaxies, we use the determination of their local 1.4 GHz luminosity function given by Sadler et al. (2002), with pure luminosity evolution of the form $(1+z)^Q$ out to $z = z_{\rm cut}$, with no further evolution thereafter (RR93); with $Q = 3.1$ (as in RR93), we find that $z_{\rm cut} = 1.5$ can reproduce the integral source counts satisfactorily, as shown in Fig. 5 (we plot integral, rather than differential, source counts as the former more clearly show the number of sources within each flux-limited sample used for the clustering analysis). For flux thresholds varying from 0.2 to 2.0 mJy, the model predicts that the starburst fraction in the integral source counts decreases from around 0.2 to 0.03. Fig. 6 shows the model redshift distributions for a flux limit of 0.2 mJy. It should be noted that the source counts (in Fig. 5) and redshift distribution (in Fig. 6, and used as an input to eqn. 5), were obtained with a cosmology of $H_0 = 50\,{\rm km\,s^{-1}\ Mpc^{-1}}$ and $\Omega_M = 1.0$, for which the above luminosity functions were also derived. However, redshift distributions and source counts are observational quantities, and thus *inde-*

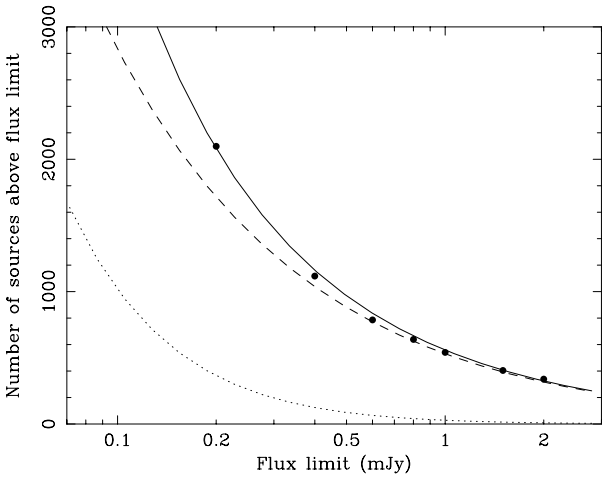


**Figure 5.** The points show the integral source counts in the central 5.32 deg$^2$ of The Bootes Deep Field where the catalogue is complete down to 0.2 mJy. The dashed and dotted lines show the contributions of the AGN and the starbursts, respectively, and the solid line their sum. The $\sqrt{N}$ error bars are smaller than the symbols. See the text for details.

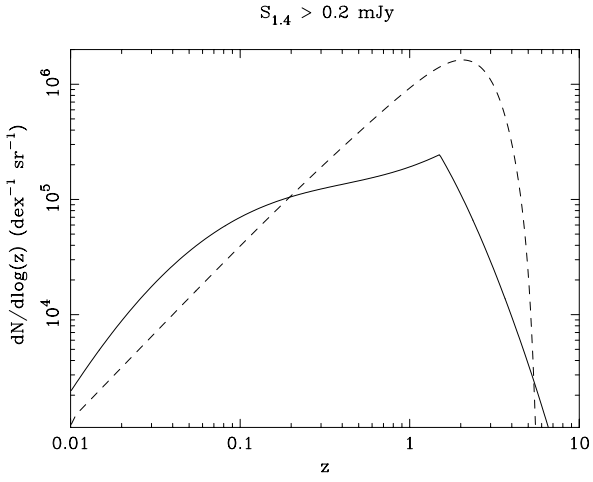


**Figure 6.** Model redshift distribution for flux limit of 0.2 mJy. The solid and dashed lines denote the starbursts and AGN, respectively.

*pendent* of the assumed cosmology. This is not, therefore, inconsistent with our use of a different cosmology ($\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$) in the clustering calculations which follow.

Using the above formalism, we compute the amplitude of the total angular correlation function for a baseline model in which $r_0 = 6.0h^{-1}$ Mpc for the AGN (comparable to that measured by BW02), and $r_0 = 3.0h^{-1}$ Mpc for the starbursts (within the range measured for local IRAS starburst galaxies). For both correlation functions, we take $\gamma = 1.8$ and the clustering evolution parameter $\epsilon = \gamma - 3$; the latter corresponds to constant clustering in co-moving coordinates, as observed for elliptical galaxies (see section 1). We take $\Omega_{\rm M} = 0.3$ and $\Omega_\Lambda = 0.7$ and note that the results are independent of $H_0$. In Fig. 7 we show, as functions of the flux limit, the amplitudes of the angular correlation functions of the AGN and starbursts separately, and for the combined population (combining the two signals as in eqn. 9). The upper limits in this figure are our 90 per cent confidence limits on the amplitude of the galaxy-galaxy clustering, obtained from a two power-law fit

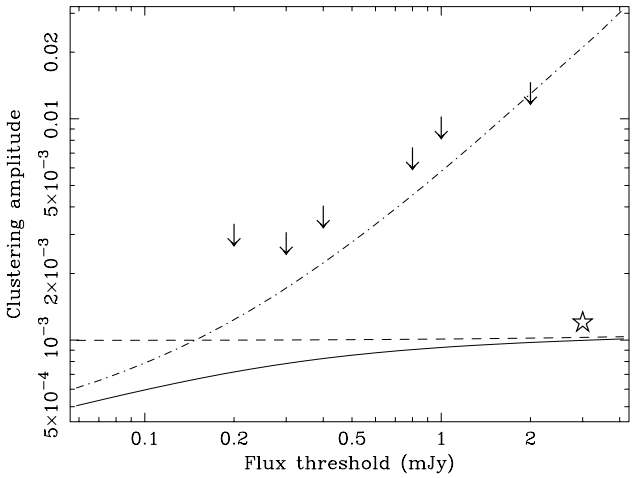


**Figure 7.** The solid line shows the model-predicted amplitude of the angular correlation function as a function of flux limit. The predicted contributions of the AGN and starbursts *alone* (i.e. not weighted by the squares of the population fractions as shown in eqn.9), are denoted by the dashed and dot-dashed lines, respectively. The arrows denote the 90 per cent confidence upper limits on the galaxy-galaxy clustering amplitude; the star is the measurement from Overzier et al. (2002) using FIRST and NVSS data.

*with the amplitude of the size distribution power-law fixed at the value set by the $1/\sigma$ extrapolation of the 5* mJy *point of BW02 (see Fig. 4).* The constancy of the model AGN clustering amplitude reflects the invariant shape of the AGN redshift distribution over this range in flux limit. The large amplitude of the starburst clustering at the high flux end reflects the fact that the brightest starbursts are local objects with a fairly narrow redshift distribution; moving towards fainter fluxes the redshift distribution broadens, reducing the amplitude. Given the small starburst fraction in the model over this flux range, the effective signal for the combined population is, with reference to eqn. (9), essentially $\omega_{\rm eff} \simeq f^2_{\rm AGN}\omega_{\rm AGN}$. There is thus little scope for measuring the clustering of starburst galaxies alone with our present data. In order to do so using unidentified radio data one would have to push down well below 0.1 mJy to the level at which the starburst fraction becomes substantially higher. However, when the $\sim 400$ IRAS-type starburst galaxies above the 0.2 mJy flux limit (see Fig. 5) can be identified in the follow-up data due on this field, measurement of their clustering strength will become easier.

## 4 CONCLUSIONS

The Bootes Deep Field is a $2.5 \times 2.5$ deg region which has been surveyed at 1.4 GHz down to an rms noise level of 28 $\mu$Jy at its centre. The area will be covered at 325 MHz, in six optical and near-infrared bands, and at longer infrared wavelengths as part of a SIRTF legacy programme. Here we have measured the angular two point correlation function of the 1.4 GHz sources down to the survey limit of 0.2 mJy. We find that the size distribution of multi-component radio galaxies dominates the overall signal even at these faint fluxes, with an amplitude consistent with the extrapolation of the $1/\sigma$ variation established over flux limits from 5 to 50 mJy by BW02 ($\sigma$ being the surface density of radio sources). This implies that the fraction of multi-component sources (normalised by the average number of components per source) varies by no more than a factor of 2 from 50 mJy to 0.2 mJy.

Only upper limits can be placed on the strength of any true galaxy-galaxy clustering. These limits are, however, consistent with the extrapolation of the amplitudes measured at higher fluxes by e.g. BW02, and with a model in which the clustering of radio-loud AGN is effectively 'diluted' by the more weakly clustered starburst galaxies. Source count models imply the latter population comprises 20 per cent of the sources above 0.2 mJy (some 400 objects), with a broad redshift distribution peaking at $z \sim 1-2$. Measurement of the clustering of these galaxies alone would extend earlier IRAS measurements from $z \sim 0.2$, but must wait until they can be identified with the follow-up data on this field.

## ACKNOWLEDGMENTS

We thank Chris Blake and Emanuele Daddi for discussions and the referees for constructive reports. The second referee is particularly thanked for bringing to our attention the numerical error in the construction of Fig. 9 of de Vries et al. (2002). RJW acknowledges support from an EU Marie Curie Fellowship.

## REFERENCES

Adelberger A., 2000, in *Clustering at high redshift*, ASP Conference Series, vol. 200, p.13
Benn C.R., Rowan-Robinson M., McMahon R.G., Broadhurst T.J., Lawrence A., 1993, MNRAS, 263, 98
Blake C., Wall J.V., 2002, MNRAS, 329, L37 (BW02)
Cress C.M., Helfand D.J., Becker R.H., Gregg M.D., White R.L., 1996, ApJ, 473, 7
Daddi E., Broadhurst T., Zamorani G., Cimatti A., Röttgering H.J.A., Renzini A., 2001, A&A, 376, 825
Daddi E., et al., 2002, A&A, 384, L1
Davis M., Peebles P.J.E., 1983, ApJ, 267, 465
de Vries W.H., Morganti R., Röttgering H.J.A., Vermeulen R., van Breugel W., Rengelink R., Jarvis M.J., 2002, AJ, 123, 1784
Dunlop J.S., Peacock J.A., 1990, MNRAS, 247, 19
Garrett M.A., 2002, A&A, 384, L19
Georgakakis A., Mobasher B., Cram L., Hopkins A., Rowan-Robinson M., 2000, A&ASS, 141, 89
Groth E.J., Peebles P.J.E., 1977, ApJ, 217, 385
Guzzo L., Strauss M., Fisher K.B., Riccardo G., Haynes M.P., 1997, ApJ, 489, 37
Hamilton A.J.S., 1993, ApJ, 419, 19
Hopkins A.M., Mobasher B., Cram L., Rowan-Robinson M., 1998, MNRAS, 296, 839
Ling E.N., Barrow J.D., Frenk C.S., 1986, MNRAS, 223, L21
Magliocchetti M., Maddox S.J., Lahav O., Wall J.V., 1998, MNRAS, 300, 257
Magliocchetti M., Maddox S.J., Lahav O., Wall J.V., 1999, MNRAS, 306, 943
McCarthy P.J., et al., 2001, ApJ, 560, L131
Moustakas L.A., Somerville R.S., 2002, in press at ApJ (astro-ph/0110584)
Oort M.J.A., 1987, PhD Thesis Leiden University
Overzier R.A., Röttgering H.J.A., Rengelink R.B., Wilman R.J., 2002, submitted to A&A
Peebles P.J.E., 1980, *The Large Scale Structure of the Universe*, Princeton
Porciani C., Giavalisco M., 2002, ApJ, 565, 24
Rengelink R.B., 1999, PhD Thesis, Leiden University
Rowan-Robinson M., Benn C.R., Lawrence A., McMahon R.G., Broadhurst T.J., 1993, MNRAS, 263, 123 (RR93)
Sadler E.M., et al., 2002, MNRAS,329, 227
Saunders W., Rowan-Robinson M., Lawrence A., 1992, MNRAS, 258, 134
Shanks T., Boyle B.J., Croom S.M., Hoyle F., Loaring N., Miller L., Outram P.J., Smith R.J., 2001, in *Mining the Sky*, Proceedings of MPA/ESO/MPE Workshop, Eds. Banday A.J., Zaroubi S., Bartelmann M., Heidelberg, Springer-Verlag
Windhorst R.A., 1984, PhD Thesis, Leiden University
Windhorst R.A., Miley G.K, Owen F.N., Kron R.G., Koo D.C., 1985, ApJ, 289, 494
Willmer C., da Costa L., Pellegrini P., 1998, AJ, 115, 869